

**По Рэйкару (model A):**

$$\ln Pr(D|\theta) = \sum_{i=1}^N \ln(a_i p_i + b_i(1 - p_i))$$

, где

- $a_i = \prod_{j=1}^R \alpha_j^{y_i^j} (1 - \alpha_j)^{1-y_i^j}$
- $b_i = \prod_{j=1}^R (\beta_j)^{1-y_i^j} (1 - \beta_j)^{y_i^j}$
- $\alpha_j = Pr[y^j = 1|y = 1]$
- $\beta_j = Pr[y^j = 0|y = 0]$
- $p_i = Pr(y = 1|x, w, H)$

$$Pr(D|\theta) = \prod_{i=1}^N (a_i p_i + b_i(1 - p_i))$$

$$E_y \ln Pr(D|\theta) = \sum_{i=1}^N \mu_i \ln(a_i p_i) + (1 - \mu_i) \ln(b_i(1 - p_i))$$

$$\mu_i = Pr[y_i = 1|y_i^j, x_i, \theta]$$

**По DS (model B):**

$$\ln Pr(D) = \sum_{j=1}^n \sum_{i=1}^m \sum_{g=1}^k [z_{ij} == g] \ln(c_{iy_j g}) = \sum_{j=1}^n \sum_{i=1}^m \ln(c_{iy_j z_{ij}})$$

, где  $i$  – работник,  $j$  – задание,  $y_j$  – настоящая метка,  $z_{ij}$  – метка работника.

По обозначениям,  $\alpha_j = c_{j11}$ ,  $(1 - \alpha_j) = c_{j10}$ ,  $\beta_j = c_{j00}$ ,  $(1 - \beta_j) = c_{j01}$ .

Таким образом,

$$Pr(D, y|\theta) = P(D|y, \theta)P(y) = \prod_{j=1}^n \prod_{i=1}^m ((\alpha_j)^{y_i^j} (1 - \alpha_j)^{1-y_i^j} P(y_j == 1))^{[y_j == 1]} ((\beta_j)^{1-y_i^j} (1 - \beta_j)^{y_i^j} P(y_j == 0))^{[y_j == 0]}$$

$$\ln Pr(D, y) = \sum_{j=1}^n \sum_{i=1}^m \ln(((\alpha_j)^{y_i^j} (1 - \alpha_j)^{1-y_i^j} P(y_j == 1))^{[y_j == 1]} ((\beta_j)^{1-y_i^j} (1 - \beta_j)^{y_i^j} P(y_j == 0))^{[y_j == 0]})$$

$$\begin{aligned} E_y \ln Pr(D, y) &= \sum_{j=1}^n \sum_{i=1}^m \mu_j \ln((\alpha_j)^{y_i^j} (1 - \alpha_j)^{1-y_i^j} P(y_j == 1)) + (1 - \mu_j) \ln((\beta_j)^{1-y_i^j} (1 - \beta_j)^{y_i^j} P(y_j == 0)) = \\ &= \sum_{j=1}^n \mu_j \ln a_j + \mu_j \ln P(y_j == 1) + (1 - \mu_j) \ln b_j + (1 - \mu_j) P(y_j == 0) = \\ &= \sum_{i=1}^N \mu_i \ln a_i + \mu_i \ln P(y_i == 1) + (1 - \mu_i) \ln b_i + (1 - \mu_i) P(y_i == 0) \end{aligned}$$

- $q_i = P(y_i == 1)$

**Смесь:**

$$\begin{aligned} P(\theta|D) &= \prod_{i=1}^N P(\theta|X_i) = \prod_{i=1}^N (P(\theta|A, X_i)P(A|X_i) + P(\theta|B, X_i)P(B|X_i)) = \\ &= \prod_{i=1}^N \left( \frac{P_A(X_i|\theta)P_A(\theta)}{P_A(X_i)} P(A|X_i) + \frac{P_B(X_i|\theta)P_B(\theta)}{P_B(X_i)} P(B|X_i) \right) \end{aligned}$$

Если  $P_A(\theta) = P_B(\theta) \sim Uni$  и  $P_A(X_i) = P_B(X_i)$ , то:

$$\begin{aligned} \theta_{MAP} &= \arg \max_{\theta} \prod_{i=1}^N \left( \frac{P_A(X_i|\theta)P_A(\theta)}{P_A(X_i)} P(A|X_i) + \frac{P_B(X_i|\theta)P_B(\theta)}{P_B(X_i)} P(B|X_i) \right) = \\ &= \arg \max_{\theta} \prod_{i=1}^N (P_A(X_i|\theta)P(A|X_i) + P_B(X_i|\theta)P(B|X_i)) \end{aligned}$$

let for any  $i$   $P(A|X_i) = \lambda$  и  $P(B|X_i) = 1 - \lambda$ .

$$\theta_{MAP} = \arg \max_{\theta} \prod_{i=1}^N (\lambda P_A(X_i|\theta) + (1 - \lambda)P_B(X_i|\theta))$$

$$\begin{aligned} P(\theta, y|D) &\propto \prod_{i=1}^N (P_A(X_i, y_i|\theta)P(A, y_i|X_i) + P(B, y_i|X_i)P_B(X_i, y_i|\theta)) = \\ &= \prod_{i=1}^N ((a_i p_i)^{[y_i==1]} (b_i (1 - p_i))^{[y_i==0]} P(A, y_i|X_i) + (a_i q_i)^{[y_i==1]} (b_i (1 - q_i))^{[y_i==0]} P(B, y_i|X_i)) = \end{aligned}$$

$$\log P(\theta, y|D) = \sum_{i=1}^N \log((a_i p_i)^{[y_i==1]} (b_i (1 - p_i))^{[y_i==0]} P(A, y_i|X_i) + (a_i q_i)^{[y_i==1]} (b_i (1 - q_i))^{[y_i==0]} P(B, y_i|X_i))$$

$$\begin{aligned} E_y \log P(\theta, y|D) &= \sum_{i=1}^N \mu_i \log(a_i p_i P(A|X_i) + a_i q_i P(B|X_i)) + (1 - \mu_i) \log(b_i (1 - p_i) P(A|X_i) + b_i (1 - q_i) P(B|X_i)) = \\ &= \sum_{i=1}^N \mu_i \log(a_i p_i \lambda + a_i q_i (1 - \lambda)) + (1 - \mu_i) \log(b_i (1 - p_i) \lambda + b_i (1 - q_i) (1 - \lambda)) = \\ &= \sum_{i=1}^N \mu_i \log a_i + \mu_i \log(p_i \lambda + q_i (1 - \lambda)) + (1 - \mu_i) \log(b_i) + (1 - \mu_i) \log((1 - p_i) \lambda + (1 - q_i) (1 - \lambda)) = \\ &= \sum_{i=1}^N \mu_i \log a_i + \mu_i \log(p_i \lambda + q_i (1 - \lambda)) + (1 - \mu_i) \log(b_i) + (1 - \mu_i) \log(1 - p_i \lambda - q_i (1 - \lambda)) \end{aligned}$$

- $p_i = \sigma(w^T x_i)$
- $q_i$  – количество единичек в массиве/априорное распределение

$$\begin{aligned}\frac{dE_y \log P(\theta, y|D)}{d\alpha^j} &= \left( \sum_{i=1}^N \mu_i (y_i^j \log \alpha^j + (1 - y_i^j) \log(1 - \alpha^j)) \right)'_{\alpha^j} = \sum_{i=1}^N \mu_i \left( \frac{y_i^j}{\alpha^j} - \frac{1 - y_i^j}{1 - \alpha^j} \right) = \\ &= \sum_{i=1}^N \mu_i \left( \frac{y_i^j - \alpha^j}{\alpha^j(1 - \alpha^j)} \right)\end{aligned}$$

$$\frac{dE_y \log P(\theta, y|D)}{d\alpha^j} = 0 \Rightarrow \alpha^j = \frac{\sum_{i=1}^N \mu_i y_i^j}{\sum_{i=1}^N \mu_i}$$

$$\beta_j = \frac{\sum_{i=1}^N (1 - \mu_i)(1 - y_i^j)}{\sum_{i=1}^N (1 - \mu_i)}$$

$$\frac{dE_y \log P(\theta, y|D)}{d\lambda} = \sum_{i=1}^N \mu_i \frac{p_i - q_i}{p_i \lambda + q_i(1 - \lambda)} + (1 - \mu_i) \frac{q_i - p_i}{1 - p_i \lambda - q_i(1 - \lambda)} = 0$$

$$\mu_i \frac{1}{p_i \lambda_i + q_i(1 - \lambda_i)} - (1 - \mu_i) \frac{1}{1 - p_i \lambda_i - q_i(1 - \lambda_i)} = 0$$

$$\mu_i(1 - p_i \lambda_i - q_i + q_i \lambda_i) - (1 - \mu_i)(p_i \lambda_i + q_i - q_i \lambda_i) = \mu_i - p_i \mu_i \lambda_i - \mu_i q_i + \mu_i q_i \lambda_i - p_i \lambda_i - q_i + q_i \lambda_i + \mu_i p_i \lambda_i + \mu_i q_i - \mu_i q_i \lambda_i = 0$$

$$\mu_i - p_i \lambda_i - q_i + q_i \lambda_i = 0$$

$$\lambda_i = \frac{q_i - \mu_i}{q_i - p_i}$$

$$\frac{dE_y \log P(\theta, y|D)}{dw_k} = \sum_{i=1}^N \mu_i \frac{\lambda_i}{p_i \lambda_i + q_i(1 - \lambda_i)} (p_i)'_{w_k} + (1 - \mu_i) \frac{-\lambda_i}{1 - p_i \lambda_i - q_i(1 - \lambda_i)} (p_i)'_{w_k}$$

$$\begin{aligned}\frac{dE_y \log P(\theta, y|D)}{d^2 w_k w_j} &= \sum_{i=1}^N \left( \mu_i \frac{\lambda_i}{p_i \lambda_i + q_i(1 - \lambda_i)} (p_i)'_{w_k} \right)'_{w_j} + ((1 - \mu_i) \frac{-\lambda_i}{1 - p_i \lambda_i - q_i(1 - \lambda_i)} (p_i)'_{w_k})'_{w_j} = \\ &= \sum_{i=1}^N \mu_i \lambda_i \frac{-1}{(p_i \lambda_i + q_i(1 - \lambda_i))^2} \lambda_i (p_i)'_{w_j} (p_i)'_{w_k} + \mu_i \lambda_i \frac{1}{p_i \lambda_i + q_i(1 - \lambda_i)} (p_i)''_{w_k, w_j} + \\ &+ (1 - \mu_i)(-\lambda_i) \frac{-1}{(1 - p_i \lambda_i - q_i(1 - \lambda_i))^2} (-\lambda_i) (p_i)'_{w_j} (p_i)'_{w_k} + (1 - \mu_i)(-\lambda_i) \frac{1}{1 - p_i \lambda_i - q_i(1 - \lambda_i)} (p_i)''_{w_k, w_j} \\ \mu_i &= \frac{a_i p_i}{a_i p_i + b_i(1 - p_i)} \lambda_i + \frac{a_i q_i}{a_i q_i + b_i(1 - q_i)} (1 - \lambda_i)\end{aligned}$$

- Если  $p_i = \sigma(w^T x_i)$ , то

$$(p_i)'_w = p_i(1 - p_i)x_i$$

$$(p_i)''_w = p_i(1 - p_i)(1 - 2p_i)x_i x_i^T$$

$$\frac{dE_y \log P(\theta, y|D)}{dw} = \sum_{i=1}^N \mu_i \frac{\lambda_i}{p_i \lambda_i + q_i(1 - \lambda_i)} p_i(1 - p_i)x_i + (1 - \mu_i) \frac{-\lambda_i}{1 - p_i \lambda_i - q_i(1 - \lambda_i)} p_i(1 - p_i)x_i$$

$$\begin{aligned} \frac{dE_y \log P(\theta, y|D)}{d^2w} &= \sum_{i=1}^N \frac{\mu_i \lambda_i p_i(1 - p_i)(1 - 2p_i)}{p_i \lambda_i + q_i(1 - \lambda_i)} x_i x_i^T - \frac{\mu_i \lambda_i^2 p_i^2(1 - p_i)^2}{(p_i \lambda_i + q_i(1 - \lambda_i))^2} x_i x_i^T - \\ &\quad - \frac{(1 - \mu_i) \lambda_i p_i(1 - p_i)(1 - 2p_i)}{1 - p_i \lambda_i - q_i(1 - \lambda_i)} x_i x_i^T - \frac{(1 - \mu_i) \lambda_i^2 p_i^2(1 - p_i)^2}{(1 - p_i \lambda_i - q_i(1 - \lambda_i))^2} x_i x_i^T \end{aligned}$$

- Если  $p_i = VGG[w]$